

Could a Classical Probability Theory Describe Quantum Systems?

Jinshan Wu^{1*}, Shouyong Pei²

1. *Department of Physics & Astronomy,
University of British Columbia,
Vancouver, B.C. Canada, V6T 1Z1*

2. *Department of Physics, Beijing Normal University
Beijing, P.R. China, 100875*

(Dated: September 12, 2008)

Quantum Mechanics (QM) is a quantum probability theory based on the density matrix. The possibility of applying classical probability theory, which is based on the probability distribution function (PDF), to describe quantum systems is investigated in this work. In a sense this is also the question about the possibility of a Hidden Variable Theory (HVT) of Quantum Mechanics. Unlike Bell's inequality, which need to be checked experimentally, here HVT is ruled out by theoretical consideration. The approach taken here is to construct explicitly the most general HVT, which agrees with all results from experiments on quantum systems (QS), and to check its validity and acceptability. Our list of experimental facts of quantum objects, which all quantum theories are required to respect, includes facts on repeat quantum measurement. We show that it plays an essential role at showing that it is very unlikely that a classical theory can successfully reproduce all QS facts, even for a single spin- $\frac{1}{2}$ object. We also examine and rule out Bell's HVT and Bohm's HVT based on the same consideration.

PACS numbers: 03.67.-a, 03.65.Yz

I. THE QUESTION AND THE COMMON GROUND TO START THE DISCUSSION

We regard quantum mechanics (QM), a theory based on wave amplitude $|\phi\rangle$ or density matrix ρ , as a quantum probability theory (QPT) as it possesses the following properties: first, for a given complete set of orthogonal vectors $\{|\mu\rangle\}$, it gives a classical probability distribution $\langle\mu|\rho|\mu\rangle$; second, for any other such vector sets related with the former one by unitary transformations, say $|\nu\rangle = \sum_{\mu} U_{\nu\mu} |\mu\rangle$, it also gives another classical probability theory, with probability distribution $\langle\nu|\rho|\nu\rangle$, which is related by the same unitary transformations, $\langle\nu|\rho|\nu\rangle = \langle\mu|U^{\dagger}\rho U|\mu\rangle$. Meanwhile, what we mean by classical probability theory (CPT), is a theory based on the classical probability distribution function (PDF) instead of the density matrix, possessing only the first property.

The goal of this work is to prove that QM could not be described by a CPT. Or put in another way, QPT could never be equivalently replaced by CPT. The question is of the existence of a map from density matrix to probability distribution function, and we want to show such map does not exist. Why we discuss this question, how we approach it, and why we claim the answer is negative will be discussed. First we clarify our terminology and establish an unambiguous language as the common starting point of this discussion.

It is necessary to distinguish between the terms QM and quantum system (QS). By QM we refer to the usual axiomatized system of quantum theory while QS is re-

served to refer to systems showing quantum properties in experiments. Second, in this work, we limit our attention to quantum measurement, excluding quantum evolution. Only axioms about quantum measurement in QM and only quantum measurement experiments are the subjects we will focus on. For example, if we say CPT can describe QS, it means CPT can explain all quantum measurement results. We wish to, for this moment, avoid the discussion of CPT on evolution of QS because evolution is less nontrivial but more technically intense. For example, we would have to construct an equivalence of Schrödinger's equation in our CPT if we wanted to discuss evolution. Furthermore, in this work, we deal only two systems, namely a single $\frac{1}{2}$ -spin system and an entangled two $\frac{1}{2}$ -spin system. In QM language, they both have finite dimension. Discussion on the above systems can easily be generalized to general QS.

The QS are systems with the following properties:

QS-I There are a set of physical quantities associated with the system whose values we can measure. For each of them, when measurement is performed on a state of the quantum system, only finite outcomes will be observed. In addition, for every single measurement, only one specific outcome appears.

QS-II If the same state is prepared, i.e. different realizations of the system go through the same preparation procedure, and the same measurement is performed on this ensemble, there is a statistical limit for the chance of appearance of a certain outcome.

QS-III If in some way the quantum system is not destroyed and it can be measured again, then

*jinshanw@physics.ubc.ca

a repeat measurement of the same physical quantity will give us the same outcome as in the last measurement, with probability 1.

QS-IV If a repeat measurement is made but of a different physical quantity, then still finite outcomes will be observed and their statistical limits also exist.

QS-V The following property is given via a specific example of a spin- $\frac{1}{2}$ system, but easily it can be generalized. Property of mixture state: the following two states can not be distinguished by any quantum measurements, including repeat measurements. Let's assume we have an apparatus preparing a spin into any desired states. Now state one is prepared as following: with probability $\frac{1}{4}/\frac{3}{4}$, we use the apparatus to prepare the spin into the *up/down*-state along z direction. State two is prepared with half possibility into the *up*-state along $\vec{r}_1 = (\theta = \frac{2\pi}{3}, \phi = \frac{\pi}{2})$ direction and half possibility into the *up*-state along $\vec{r}_2 = (\theta = \frac{2\pi}{3}, \phi = \frac{3\pi}{2})$ direction.

Usual QM realize those properties of a QS through axioms:

QM-I States of a quantum object are normalized vectors in a complex linear space \mathcal{H} with dimension N , equipped with a definition of inner product. Or equivalently, state of this object is described by a $N \times N$ hermitian positive-defined normalized matrix ρ . The set of such density matrices is denoted as $\mathcal{N}(\mathcal{H})$, the normalized positive operators over \mathcal{H} .

QM-II Physical quantities are hermitian operators over \mathcal{H} . Their set is denoted as $\mathcal{O}(\mathcal{H})$. Physical quantities are measurable. The measurement of A on a system at state ρ , results event α (meaning value of observable A is recorded as α) with probability p_α . α is one of the eigenvalues of A (assumed non-degenerate but could be trivially generalized) and $p_\alpha = \langle \alpha | \rho | \alpha \rangle$.

QM-III The state of the object after measurement, given the observed value is α , is $|\alpha\rangle\langle\alpha|$.

In a finite dimensional Hilbert space, the number of eigenvalues of an operator is finite. QM-II realizes both QS-I and QS-II. QM-III realizes QS-III in that if the same measurement is repeated, the outcome must be α and with probability 1.

In order to realize QS-IV, one needs to consider basis transformations in Hilbert space, that one vector could be expanded under difference bases. After measurement of A provided the outcome is event α , the system stays at $\rho = |\alpha\rangle\langle\alpha|$. If one then measures for example B with eigenvalues $\{\beta\}$, then according to QM-II, the event of a

specific β will appear with probability $p_\beta = \langle \beta | \rho | \beta \rangle = \langle \beta | \alpha \rangle \langle \alpha | \beta \rangle$. This explain QS-IV.

For, in QS-V, in usual QM language both preparations result the same mixture state, $\rho = \frac{1}{4}|\uparrow\rangle\langle\uparrow| + \frac{3}{4}|\downarrow\rangle\langle\downarrow|$. Therefore no measurement can tell their difference. And conceptually we want our alternative theory, whatever it is, to respect the idea that the mixture state, as in experiment of QS-V, is a probability summation of all the exclusive possibilities. Our usual QM does so. For example, we can confirm state two leads to the same density matrix via the probability summation rule,

$$\begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{3}{4} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{1}{4} & i\frac{\sqrt{3}}{4} \\ -i\frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} \frac{1}{4} & -i\frac{\sqrt{3}}{4} \\ i\frac{\sqrt{3}}{4} & \frac{3}{4} \end{bmatrix}. \quad (1)$$

Therefore QM realizes all the five experimental facts of a QS.

Next we construct a classical theory for all five experimental facts. Such theory does not need to respect the QM axioms at all, but it should still respect QS-I, QS-II, QS-III, QS-IV and QS-V. In a usual discussion of HVT, only the first two are required to be respected by the theory. We will see that if only these two are required it is not impossible to have a classical theory.

We have to mention that as an experimental fact, QS-III/IV is not unquestionable. Usually the state is destroyed after measurement. However, quantum non-demolition measurements[4] (QNM) allows a system to be subject to repeat measurements. Therefore we take QS-III/IV also as an experimental fact. Another thing worth mentioning is we did not include finite accuracy of real measurements into our experimental facts. Our usual QM (QM-I, II, III) embraces non-zero commutators between operators so it support the idea of Uncertainty Principle. However, as argued by Bohm [1], on the fundamental level one could not tell if it is really impossible to measure some quantities simultaneously or it is just because of problems on technology or accuracy of experiments. This gives us the possibility to relax non-commutation relations between physical quantities when necessary.

The question of the possibility of fulfilling QS-I, II, III, IV and V by a CPT, requires to solidify explicitly what we refer to as a CPT. It means the following:

CPT-I States form a set of event Ω . There is a map P from σ -Algebra \mathcal{F} of Ω to $[0, 1]$. P satisfies the Kolmogorov axioms of probability[2]. Only physical quantities corresponding to members of \mathcal{F} are observable. A simpler case, which is quite often the case of a physical system, is that the set Ω is a set of countable simple events and \mathcal{F} is the trivial topology, set of all subsets of Ω . In our discussion, we only work with this simpler case. For exclusive events, if $A \cap B = \phi$, $A, B \in \mathcal{F}$, then $P(A \cup B) = P(A) + P(B)$. And for independent events, $A \otimes B \in \mathcal{F}(\Omega_1 \otimes \Omega_2)$ where

$A \in \mathcal{F}(\Omega_1), B \in \mathcal{F}(\Omega_2)$, then $P(A \otimes B) = P(A) \cdot P(B)$.

CPT-II When the measurement of any $A \in \mathcal{F}$ is performed, every value of $\omega \in A \subseteq \Omega$ could be observed, with corresponding probability $P(\omega)$.

CPT-III After the measurement, the state of the system is the one observed. Provided event ω is recorded, the state of the object after measurement is ω .

The validity of CPT-III is not really explicitly defined in the usual probability theory. CPT itself provides no answer at all for that, but usually people like to interpret CPT in this way. For example, imagine a truly random perfect dice. After it is measured, one would like to say it is at the state just observed by us. However, experimentally what it really means is if the dice is measured again, it guarantees one will observe the same value with probability 1. Therefore, although CPT-III is quite natural, it can be altered if necessary.

After clarifying the terminology, now the question we seek to discuss is better defined. We are looking for a CPT, which follows CPT-I, CPT-II, CPT-III, of two systems: a single $\frac{1}{2}$ -spin and two entangled $\frac{1}{2}$ -spins, which both possess QS-I, QS-II, QS-III, QS-IV and QS-V. We will set the state of the single-spin system at the *up*-state along the x direction and the two-spin system at a singlet state. Although we aim at using CPT as an alternative theory for quantum systems, we will still use the usual QM language to denote their states. In another words, we admit QM is a theory for quantum systems but we seek to determine if QS can also be described by a classical theory such as CPT.

In section §II we will put both CPT and QPT into a density matrix form so that we can use the same mathematical language to discuss the two theories. In section §III we discuss why we want to have such a map. Afterwards we will present a CPT for a single-spin system and a CPT for a two-spin system, in section §V and §VI respectively. We will see that what kind of CPT is necessary to fully describe quantum systems. We will then discuss why our CPT violates Bell's inequality and what is the possible interpretation of such CPT. Finally in section §VII we conclude that if we are willing to accept all the prices we have to pay to have such a CPT for quantum system, our CPT could be the one. But it is even harder to be understood as compared with the usual QM.

II. DENSITY MATRIX LANGUAGE FOR BOTH CLASSICAL AND QUANTUM SYSTEMS

In density matrix language for QM, the state of a quantum object is represented by a density matrix $\rho^q(t)$. The evolution is described by a unitary transformation

$U(t) \triangleq U(0, t)$ as

$$\rho^q(t) = U(t) \rho^q(0) U^\dagger(t), \quad (2)$$

where generally $U(t)$ is determined by H , the Hamiltonian of the quantum object. For a pure initial state, the above density matrix formalism is equivalent with the usual wave function or right vector formalism, but it can also describe a mixture state. For example, we can consider an exclusive mixture state as used in Von Neumann's picture of quantum measurement[3],

$$\rho^q = \sum_i p_i |\phi_i\rangle \langle \phi_i|, \quad (3)$$

where $\{|\phi_i\rangle\}$ is a set of orthogonal normalized vectors. According to Von Neumann's picture, the meaning of such an exclusive state is that every sample of this object chooses one of $\{|\phi_i\rangle\}$ with probability p_i .

This explanation reminds us of the PDF of a truly random classical object (TRCO), which generally should be included as objects of classical mechanics (CM). A state of a TRCO is a PDF $p(x)$ normalized over $\Omega = \{x\}$, the set of all its possible states. It's therefore possible to rewrite this PDF as a density matrix

$$\rho^c = \sum_{x \in \Omega} p(x) |x\rangle \langle x|, \quad (4)$$

which gives exactly the same information provided by a PDF. One could just regard this as another notation of a discrete PDF. For the purpose of normalization we also require that simple events are exclusive,

$$\langle x | x' \rangle = \delta(x - x'), \quad (5)$$

where $\delta(x - x')$ is the Kronecker delta for the discrete set Ω . In QM, generally a state of a quantum system is a full structure density matrix, while in CM, a state of a TRCO is a diagonal density matrix. Considering only discrete sets allow us to use the notation $|x\rangle$ and inner product $\langle x | y \rangle = \delta_{xy}$ without any problem. Although physicists also use such notation for continuous systems, mathematicians do not like the idea of using $|x\rangle$ as a basis vector, or using Dirac δ function as a basis of function space, for continuous system. Most expressions in physicists' notation can be mapped onto more rigorous mathematicians' notation[5], but we don't wish to deal with that here. We limit our description to discrete systems only. For example, a perfect dice is such an object.

From this point forward, we are going to use density matrix notation for both QS and TRCO. Furthermore, we can construct a similar theory to describe classical evolution processes. For example, if we denote the evolution process as a linear operator \mathcal{T} , then time evolution of such classical objects can be defined as

$$\rho^c(t) \triangleq \mathcal{T}(\rho^c(0)) = \sum_x p(x) \mathcal{T}(|x\rangle \langle x|). \quad (6)$$

Formally, we can use the evolution operator T as $\mathcal{T}(|x\rangle\langle x|) = (T|x\rangle)(\langle x|T^\dagger) = |x(t)\rangle\langle x(t)|$, so

$$\rho^c(t) = T\rho^c(0)T^\dagger. \quad (7)$$

Also $TT^\dagger = T^\dagger T = I$, which can be proved as follows: first, for system fully determined by x , we have $\delta(x(t) - y(t)) = \delta(x - y) = \langle x|y\rangle$, then

$$\langle x|T^\dagger T|y\rangle = \langle x(t)|y(t)\rangle = \delta(x(t) - y(t)). \quad (8)$$

Therefore, both QM and CPT are unitary evolution theories of density matrices, while the difference between them is the existence of off-diagonal elements. From this point of view, our task in this work is to put a full-structure density matrix into a diagonal density matrix. We call this a question of finding a diagonalization map. The reason we introduce TRCOs is to help towards the understanding of classical objects and, later, quantum objects. By emphasizing “truly random”, we are not referring to objects which behave randomly because of the uncertainty in their initial conditions. For a TRCO there is intrinsically no way, even for “God”, to tell its real state before a measurement is performed. We only can say it stays in a classical mixture state. One may argue that a physical classical object is not a truly random object. Imagining such a TRCO, however, will help us to understand the classical and quantum measurement process.

To conclude this section, we want to point out that our language of the diagonal density matrix for CPT and the non-diagonal density matrix for QM provides a unified description of classical and quantum mechanics. Besides this, there is another set of language based on C^* -algebra[6] also unifying the description of classical and quantum mechanics. There, classical and quantum operators are more basic descriptions of a system and they form an abelian C^* -algebra and a non-abelian C^* -algebra respectively. States are defined as functionals over the corresponding algebra. Although we will not prove it explicitly here, we believe that our notation of the diagonal and non-diagonal density matrix for classical and quantum systems, is in fact equivalent with the C^* -algebra based language.

III. WHY ARE WE LOOKING FOR SUCH A MAP?

If we have a CPT as desired, it is a HVT of quantum systems. Quantum systems are no longer quantum but TRCOs. Therefore, one can understand quantum measurement if one can understand measurement of TRCOs. The most straightforward picture of a measurement is a measurement on a determinant classical object. Assumed as a discrete system, it stays in state $|x\rangle\langle x|$ before it is measured. After the measurement we get the information that it was in state $|x\rangle\langle x|$ and it remains in state $|x\rangle\langle x|$. The less straightforward picture of a measurement is a measurement on a statistically random classical

object. Here the term “statistically random” means that the nature of this object is still determinant, but with incomplete information it appears as a random object. For every given such object, we just do not know its state but it is already fixed. This also means its randomness is only meaningful as in an ensemble. This is called statistical interpretation of probability theory. Again for such an object, it is in state $|x\rangle\langle x|$ before measurement and after the measurement we get the information that it was in state $|x\rangle\langle x|$ and it remains in state $|x\rangle\langle x|$. Notice that although x can be one of a large set, but it is fixed with probability $p(x)$ before the measurement is performed.

Measurement on a TRCO is less understandable. Assuming such an object really exists for the moment, its state is unknown before measurement. After the measurement we find with probability $p(x)$ that it was in state $|x\rangle\langle x|$ and it remains in state $|x\rangle\langle x|$ afterwards. Here we find that a phenomena so-called “collapse” of probability function has occurred. While this seems less understandable, both “statistical randomness” and “true randomness” give us the same measurement result. One could never distinguish which is the “real” one from measurements. It is a philosophical question to ask which one it really is, statistically random or truly random and so from now on we will treat them as the same.

Now imagine we have two correlated TRCOs which have exactly the same states, but unknown. Since they are both TRCOs we do not know their states before the measurement. If we measure one of them, say we find that it stays in state $|x\rangle\langle x|$, then we immediately know the state of the other object is also $|x\rangle\langle x|$. In this sense, if we assume the existence of such TRCOs, “spooky action” exists even in classical mechanics. A classical bit of information need to be transferred from one to the other in order for the other to know that its counterpart’s state after the measurement. Quantum “spooky action” in entangled systems is not stranger than its classical version at all. They are different just that in the quantum case, both direction and measurement outcome need to be transferred, not only the outcome.

Provided there is a TRCO description of quantum systems, the two problems of quantum measurement, namely collapse of the wave function and measurement of entangled states, become collapse of the probability function and measurement of classical correlated states in measurement of TRCOs. This implies that if one believes measurement of TRCOs is understandable, then measurement of quantum systems is also understandable.

Here we assume TRCO Assumption: there is no difficulty or confusion in understanding measurement of TRCOs. Even if it is questionable, if TRCO can describe quantum systems, then we know the problem of quantum measurement comes from classical probability theory and has nothing to do with any other quantum nature. Of course the situation will be different if we find out that TRCOs can not describe quantum systems.

We can formally compare measurement of TRCOs and quantum systems. Here we include both auxiliary system

m and object system o explicitly into our formal description. The measurement includes two steps. First, a classical correlated state is formed by an interaction process, so that from an initial state

$$\rho^{c,o} = \sum_x p(x) |x\rangle \langle x|. \quad (9)$$

we get

$$\rho^{c,o} \otimes \rho^{c,m} \longrightarrow \rho^{c,om} = \sum_x p(x) |x \otimes M(x)\rangle \langle x \otimes M(x)|. \quad (10)$$

Second, when we only check the value recorded on the auxiliary system, we get a sample from the auxiliary system's partial distribution, which is

$$\rho^{c,m} \triangleq \text{tr}^o(\rho^{c,om}) = \sum_x p(x) |M(x)\rangle \langle M(x)|, \quad (11)$$

where tr^o means the trace is taken over object state space, a standard procedure in probability theory when only information on the partial distribution is needed. Therefore, according to their exclusiveness nature and CPT-III, the sampling process gives us one specific state $M(x^*)$. This happens with the desired probability $p(x^*)$, due to CPT-II. $M(x^*)$ on the auxiliary system means x^* on the measured object.

However, even formulated in the same way but in usual QM language, the picture of quantum measurement is different because the general quantum density matrix has non-zero off-diagonal terms. As in equ(10) and equ(11), with first an interacting process and then a partial trace, if the same steps are applied onto a quantum system with

$$\rho^{q,o} = \sum_{\mu\nu} \rho_{\mu\nu} |\mu\rangle \langle \nu|, \quad (12)$$

then, firstly,

$$\rho^{q,o} \otimes \rho^{q,m} \longrightarrow \rho^{q,om} = \sum_{\mu\nu} \rho_{\mu\nu} |\mu \otimes M(\mu)\rangle \langle \nu \otimes M(\nu)|, \quad (13)$$

and secondly, when we only check the value recorded on the auxiliary system, we obtain a sample from the auxiliary system's partial distribution, which is a sample of

$$\rho^{q,m} \triangleq \text{tr}^o(\rho^{q,om}) = \sum_{\lambda} \rho_{\lambda\lambda} |M(\lambda)\rangle \langle M(\lambda)|. \quad (14)$$

However, for a quantum object, equ(14) is not a copy of equ(12), while equ(11) is an exact copy of equ(9) for a classical object. Therefore, if TRCOs could never describe quantum system, even with the TRCO Assumption, quantum measurement is still harder to understand than measurement of TRCOs. However, if we have a CPT for quantum system, then quantum measurement is just as understandable as measurement of a TRCO.

As we have seen, due to CPT-III, a classical measurement ends up with an exact copy of the object state.

We may regard such a process as a clone. However, this clone does not respect the definition of clone in the original quantum non-cloning theorem[8] (QNCT),

$$\rho_{aim}^o \otimes \rho_{initial}^m \xrightarrow{U} \rho^{om} = \rho_{aim}^o \otimes \rho_{aim}^m, \quad (15)$$

while now it has more general property $\rho_{aim}^o \otimes \rho_{initial}^m \xrightarrow{U} \rho^{om}$ that,

$$\text{tr}^m(\rho^{om}) = \rho_{aim}^o \text{ and } \text{tr}^o(\rho^{om}) = \rho_{aim}^m, \quad (16)$$

Equ(15) is a special case of equ(16). In fact, this more general clone is called a broadcast and it has been proved that a quantum system can not be broadcasted in quantum no-broadcasting theorem (QNBT)[9]. Unless the object system initially stays in one of a set of **known** orthogonal states, a quantum system can not be broadcasted. In our language, this means when a system is in a classical probability combination of known orthogonal states, i.e a diagonal density matrix under a known basis, it can be broadcasted. This is just a broadcast of TRCOs.

An arbitrary unknown state of a TRCO can be broadcasted, or a diagonal density matrix state can be broadcasted. Therefore, if the above diagonalization mapping exists, through it, a quantum system can also be broadcasted. This would conflict with QNBT, which is proved in the language of usual QM. This leads to two possibilities: firstly, QNBT holds and diagonalization mapping does not exist; or secondly, QNBT is not valid and the mapping exists. Now we find that QNBT is also reduced to the existence of the diagonalization mapping. Therefore, it seems all the confusing and "extraordinary" problems in QM including quantum measurement, HVT and QNBT come down to one question, the existence of such diagonalization mapping.

The relation between QNBT and HVT can be shown more explicitly. A TRCO can be broadcasted, by introducing a classical hidden variable. For example, let us use a perfect two-face dice as a TRCO. We introduce a classical signal λ , generated from a given PDF $\rho(\lambda)$ over $\Gamma = \{\lambda\}$. The state of the dice is determined by this signal as follows,

$$\rho^{c,o} = \sum_{\lambda \in \Gamma} \rho_+(\lambda) |+\rangle \langle +| + \sum_{\lambda \in \Gamma} \rho_-(\lambda) |-\rangle \langle -|, \quad (17)$$

where it is required that

$$\int_{\Gamma} d\lambda \rho_+(\lambda) = \frac{1}{2} = \int_{\Gamma} d\lambda \rho_-(\lambda). \quad (18)$$

We then duplicate this hidden variable signal, send a copy to another dice while the original signal is sent to the original dice. Each dice determines its state respectively according to the value of its hidden variable. Now we get a broadcast of the dice. In this sense, it is fair enough to say that the success of an HVT for CM makes it possible to broadcast a classical object. So what about a HVT for QM?

IV. A POSSIBLE TRCO AND UNDERSTANDING OF ITS MEASUREMENT

Consider a quantum system coupled with a large thermal bath whose eigenenergy can be measured in much shorter time than the relaxation time. Our measurement is performed once in a while with the time interval between measurements being much longer than the typical relaxation time of this system. The outcomes of such measurements will give us a sequence of eigenvalues whose probability of appearance follows classical Boltzmann distribution. Do we now believe that the system stays in one of the eigenstates before any measurements? And further, does our belief matter? It seems there is no difficulty in accepting the results from this measurement as is. From this example, we wish to argue that our assumption of the existence of TRCO and validity of TRCO Assumption, which states there is no problem in understanding measurement of TRCOs, is plausible.

V. CPT FOR SINGLE-SPIN SYSTEM

The possibility of a CPT or a HVT for quantum system has been long investigated by many great physicists[1, 10, 11, 12, 15]. Bell's Theorem[10] says that all local HVT should obey the Bell's inequality, which is not respected by QM. Experimental tests suggests that QS do violate the Bell's inequality so QM is a preferred theory for QS[14]. But this statement has not yet been supported by all physicists. In the following, we will try to answer this problem in another way. We are willing to go as far as possible to construct a CPT to give consistent results with quantum systems including QS-I, II, III, IV and V. If this effort fails we will find where and why; or if it succeeds, we will check whether it is acceptable or not. If it succeeds, according to Bell's Theorem, it should be non-local. It will be interesting to show explicitly the place where non-locality enters the theory. In fact, in [15], the author already discussed a similar question of "How to make quantum mechanics look like a hidden-variable theory and vice versa" using the Wigner distribution. Here in this paper, to discuss the same question, we start from a more general form of CPT and try to make it successful as far as possible. For simplicity of language, in this paper, we regard CPT and HVT of a quantum system as being the same meaning and later on just simply call them HVT.

According to our general framework, HVT could be in a classical diagonal density matrix form,

$$\rho^{hvt} = \sum_{\lambda \in \Gamma} \rho(x(\lambda)) |x(\lambda)\rangle \langle x(\lambda)|, \quad (19)$$

where x is the dynamical variable, λ is the hidden random variable and $x(\lambda)$ is an onto mapping, $\rho(x(\lambda))$ is a PDF over $\Gamma = \{\lambda\}$, a set of exclusive events,

$$\langle x(\lambda) | x(\lambda') \rangle = \delta(\lambda - \lambda'). \quad (20)$$

One thing that is necessary to be pointed out is here the parameter λ is abstract, not limited as a single variable. A successful HVT has to respect all QS facts. We will start from QS-I and II.

A. CPT based on exclusiveness of all elementary pure events

We first consider a single spin- $\frac{1}{2}$ as in Bohm's HVT[1], and then focus on an entangled object with two subsystems as discussed in Bell's inequality[10]. For simplicity, let's just consider a specific quantum state, a $\frac{1}{2}$ -spin in the state of $|\uparrow\rangle_x$, the *up* state of S_x . In the language of QM, it's

$$\rho^q = \frac{1}{2} (|\uparrow\rangle_z \langle\uparrow|_z + |\uparrow\rangle_z \langle\downarrow|_z + |\downarrow\rangle_z \langle\uparrow|_z + |\downarrow\rangle_z \langle\downarrow|_z). \quad (21)$$

For a HVT, the first trial density matrix will naturally be,

$$\rho^{hvt} = \sum_{\lambda_z} \rho_+(\lambda_z) |\uparrow\rangle_z \langle\uparrow|_z + \sum_{\lambda_z} \rho_-(\lambda_z) |\downarrow\rangle_z \langle\downarrow|_z, \quad (22)$$

with the following requirement to give correct results for measurement on S_z ,

$$\int_{\Gamma_z} d\lambda_z \rho_+(\lambda_z) = \frac{1}{2} = \int_{\Gamma_z} d\lambda_z \rho_-(\lambda_z). \quad (23)$$

However, this gives the consistent results with QS-I and II only for S_z measurement. We can also measure S_x . If we still respect the possible non-commutative relation between quantum operators S_x and S_z , then we need to do a basis transformation in \mathcal{H}^q and do measurement of S_x . We get

$$\begin{aligned} \rho^{hvt} = & \frac{1}{2} \sum_{\Gamma_z} [\rho_+(\lambda_z) + \rho_-(\lambda_z)] (|\uparrow\rangle_x \langle\uparrow|_x + |\downarrow\rangle_x \langle\downarrow|_x) \\ & + \frac{1}{2} \sum_{\Gamma_z} [\rho_+(\lambda_z) - \rho_-(\lambda_z)] (|\uparrow\rangle_x \langle\downarrow|_x + |\downarrow\rangle_x \langle\uparrow|_x). \end{aligned} \quad (24)$$

We can see that, according to equ(23), the result of this measurement will be $\frac{1}{2}$ probability to get *up* and $\frac{1}{2}$ to get *down*. This is obviously wrong. We know for the specific state we choose above, the correct result of the S_x measurement is the *up* state only. This HVT does not realize QS-I and II.

There is one way to overcome this inconsistency with the price that not one hidden variable, but another hidden variable is needed. In order to get correct results for measurement on S_z and S_x , we need

$$\begin{aligned} \rho^{hvt} = & \frac{1}{\mathcal{N}} \left[\sum_{\lambda_z} \rho_+(\lambda_z) |\uparrow\rangle_z \langle\uparrow|_z + \sum_{\lambda_z} \rho_-(\lambda_z) |\downarrow\rangle_z \langle\downarrow|_z \right. \\ & \left. + \sum_{\lambda_x} \rho_+(\lambda_x) |\uparrow\rangle_x \langle\uparrow|_x + \sum_{\lambda_x} \rho_-(\lambda_x) |\downarrow\rangle_x \langle\downarrow|_x \right], \end{aligned} \quad (25)$$

with the requirement,

$$\int_{\Gamma_x} d\lambda_x \rho_+(\lambda_x) = 1, \int_{\Gamma_x} d\lambda_x \rho_-(\lambda_x) = 0. \quad (26)$$

\mathcal{N} is a normalization constant to keep $\text{tr}(\rho) = 1$ and here $\mathcal{N} = 2$. With this density matrix, a measurement of S_x will give the *up* state only. We can similarly include S_y terms using another hidden variable λ_y . However, a successful HVT should respect QS-I and II for measurement on an arbitrary direction. For this purpose, will three hidden variables corresponding to S_x, S_y, S_z be enough? For example, for a measurement of

$$S_r = \sin \theta \cos \phi S_x + \sin \theta \sin \phi S_y + \cos \theta S_z, \quad (27)$$

on the above state, the possible outcomes are

$$s_r = \frac{1}{2} (\sin \theta \cos \phi \pm \sin \theta \sin \phi \pm \cos \theta). \quad (28)$$

This could be continuous number, not only $\pm \frac{1}{2}$. We see that it does not respect QS-I. So QS-I requires one hidden variable for measurement on every direction and abandonment of the inherent relation between operators such as equ(27). Furthermore such multi-hidden variable density matrix has one very important implication, that according to equ(20), all states (events) corresponding to arbitrary directions should all be exclusive events. This implies $\sigma_{\vec{r}_1} \sigma_{\vec{r}_2} = 0$ and our CPT density matrix has to be

$$\rho^{hvt} = \frac{1}{\mathcal{N}} \sum_{\vec{r}} [p_{\uparrow}(\vec{r}) |\uparrow\rangle_{\vec{r}} \langle \uparrow|_{\vec{r}} + p_{\downarrow}(\vec{r}) |\downarrow\rangle_{\vec{r}} \langle \downarrow|_{\vec{r}}] \quad (29)$$

where

$$p_{\uparrow}(\vec{r}) = \frac{1+r_x}{2}, p_{\downarrow}(\vec{r}) = \frac{1-r_x}{2}. \quad (30)$$

There is a technical problem and another non-trivial conceptual problem with the above PDF. The technical problem is the value of \mathcal{N} . Since we need to keep $\text{tr}(\rho^{hvt}) = 1$ and there is infinite number of directions, \mathcal{N} will be infinity if $\text{tr}(\rho^{hvt})$ is simply,

$$\text{tr}(\rho^{hvt}) = \sum_{\vec{r}} [\langle \uparrow|_{\vec{r}} \rho^{hvt} |\uparrow\rangle_{\vec{r}} + \langle \downarrow|_{\vec{r}} \rho^{hvt} |\downarrow\rangle_{\vec{r}}]. \quad (31)$$

One way to define a “proper” $\text{tr}(\rho^{hvt})$ to avoid such divergence is to decompose $\vec{r} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ and treat

$$\text{tr}(\rho^{hvt}) = \int d\theta d\phi \sin \theta [\langle \uparrow|_{\vec{r}} \rho^{hvt} |\uparrow\rangle_{\vec{r}} + \langle \downarrow|_{\vec{r}} \rho^{hvt} |\downarrow\rangle_{\vec{r}}]. \quad (32)$$

In this case, $\mathcal{N} = 4\pi$. This introduces additional relative probability between states corresponding to different \vec{r} . This may not be a proper definition, however, it is still possible to solve this technical question of divergent normalization constant by some other ways. If only relative

probability of a given direction \vec{r} is concerned in real measurements, this problem does not affect the outcomes at all.

The other problem is rather serious. That is due to the full exclusiveness between all events, the meaning of a measurement changes. “Measuring $\sigma_{\vec{r}}$ ” for a specifically given \vec{r} is no longer a pure elementary event but a compound event. A pure elementary event instead would be “measuring σ ”, with no specific direction given. The result of such a measurement will be one direction, which got randomly picked up during the measurement process, and an *up*- or *down*- state, would be recorded correspondingly with the right probability. In this way, there is no guarantee that the randomly picked-up direction will be the desired direction of an observer.

A classical dice would be a good example of a classical probability distribution based on all exclusive events. From a perfect 6-face dice, we wish to only measure the relative probability between face 1 and face 2. We could still get all 6 numbers, but we discard all the other four if they turn out to be the outcomes of our measurement. Therefore, effectively we will find out the state of the dice within the subspace is,

$$\rho^c = \frac{1}{2} (|1\rangle \langle 1| + |2\rangle \langle 2|). \quad (33)$$

Similarly measurement of σ on our ρ^{hvt} will be one of all of the exclusive events, and during our analysis of the results, we can discard all irrelevant events. In real quantum measurements, however, we never find such irrelevant and redundant outcomes. If we measure σ_x , according to the all exclusive nature, in a classical measurement of the above state, sometimes our apparatus detects nothing and sometimes it detects the right state – *up*. However, in a real quantum measurement, assuming no further experimental accuracy limits, such detecting-nothing events never happen. What we get is only the *up*- or *down*- state of a given direction. This shows that in fact the above state based on exclusiveness is not the desired state. Or if it is then this is only possible if the system somehow knows the intention of the observer during the process.

This “contextual” relation between system and observer is unexpected, however, some physicists may still be willing to accept such a theory since it is a problem about interpretation of a theory not about any predictions from the theory.

Now we will try to make this HVT compatible with QS-III and QS-IV, those facts about repeat measurement. CPT-III tells us that if all events are exclusive, then after a measurement, for example along the x direction, given the *up* state is recorded, its state is simply $|\uparrow\rangle_x \langle \uparrow|_x$. A repeat measurement along the x direction results in event *up* again. This is the expected result stated in QS-III. However, if the repeat measurement is along the z direction, for state $|\uparrow\rangle_x \langle \uparrow|_x$, the exclusiveness tells us, there is not any such events as measurement along z direction. So we would again get a detecting-nothing

event. This conflicts with QS-IV. Furthermore, the initial state is the *up* state along the x direction, therefore after a measurement along the x direction, nothing changes. If CPT-III holds, we see the state before and after the measurement is respectively,

$$\rho_{before}^{hvt} = \sum_{\vec{r}} [p_{\uparrow}(\vec{r}) |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + p_{\downarrow}(\vec{r}) |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}] \quad (34)$$

and

$$\rho_{after}^{hvt} = |\uparrow\rangle_x \langle\uparrow|_x, \quad (35)$$

where

$$p_{\uparrow}(\vec{r}) = \frac{1+r_x}{2}, p_{\downarrow}(\vec{r}) = \frac{1-r_x}{2}. \quad (36)$$

Equ(34) and equ(35) are obviously different. The state stays the same before and after the measurement, however, we find their expressions are different. This means CPT-III is wrong. A state after it was revealed in a measurement is not the state corresponding to the measurement result. We will have to also sacrifice CPT-III after abandoning equ(27).

CPT-III' After a measurement, the object stays at the state which guarantees a repeat measurement in accordance with QS-III and QS-IV. For example, for spin- $\frac{1}{2}$ after a measurement on $\sigma_{\vec{r}_0}$ and a *up*-state being recorded, the state is,

$$\rho_{after}^{hvt} = \sum_{\vec{r}} \left[\frac{1+\vec{r} \cdot \vec{r}_0}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{1-\vec{r} \cdot \vec{r}_0}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}} \right]. \quad (37)$$

If a *down*-state is recorded after measurement of $\sigma_{\vec{r}_0}$, we can simply replace \vec{r}_0 with $-\vec{r}_0$ in equ(37).

This CPT-III' is not easily understood.

Furthermore, this is not a consistent theory. We already know that an *up*-state along the x direction, before and after measurements of σ_x , is equ(34). Given this state if we want to calculate the probability of observing the *up*-state along the x direction, we will do

$$\begin{cases} p_{x_{up}} = \langle\uparrow|_x \rho_{before}^{hvt} |\uparrow\rangle_x = \frac{1}{\mathcal{N}}, \\ p_{x_{down}} = \langle\downarrow|_x \rho_{before}^{hvt} |\downarrow\rangle_x = 0. \end{cases} \quad (38)$$

This gives the correct answer that the relative probability between *up*- and *down*-state is 1. But notice that we times $\langle\uparrow|_x$ from the left and $|\uparrow\rangle_x$ the right to a density matrix to get the probability of $p_{x_{up}}$. In doing so we assume that vector $|\uparrow\rangle_x$ stands for the event of an *up*-state along the x direction, but it is different from equ(34), which is the expression standing for the event of an *up*-state along the x direction as we pointed out before. We have now two different expressions for the same state in a theory.

Therefore we conclude that the first rescue of HVT, based on the assumption of the exclusiveness among all $\{\sigma_{\vec{r}}\}$, failed to achieve a consistent theory satisfying simultaneously QS-I, QS-II, QS-III and QS-IV. To do so we will not have inherent relation between operators as in equ(27), we will have to put a twist on CPT-III and allow “contextual” communication between object and observer. Even after all these, we would not be able to get a self-consistent theory. We will now try out another more plausible construction of CPT for QS, based on independence of all pure elementary events.

B. CPT based on independence of pure elementary events

Although the idea of exclusive events fails, in fact, there is another way to save the idea of HVT, being that λ_z and λ_x are independent events, so that a HVT density matrix could be,

$$\rho^{hvt} = \sum_{\{\vec{\lambda}\}} \rho(\vec{\lambda}) |\dots, x_{\vec{r}}(\lambda_{\vec{r}}), \dots\rangle \langle\dots, x_{\vec{r}}(\lambda_{\vec{r}}), \dots|, \quad (39)$$

where $\lambda_{\vec{r}}$ is a random variable for direction \vec{r} and $x_{\vec{r}}(\lambda_{\vec{r}}) = \uparrow, \downarrow$. Notation $\vec{\lambda}$ refers to an infinite dimensional vector $(\lambda_x, \dots, \lambda_y, \dots, \lambda_z, \dots)$. Under this independent event assumption, measurement on every direction is done independently. This is only possible if all $\sigma_{\lambda_{\vec{r}}}$, operators corresponding to all directions $\lambda_{\vec{r}}$ are commutative and every operator could be treated independently.

HVT requires again abandoning inherent relation as in equ(27) and non-commutative relation between quantum operators. A valid multiplication between operators is the direct product, $\sigma_{\vec{r}_1} \otimes \sigma_{\vec{r}_2}$. A common basis is $|\uparrow_x \text{ or } \downarrow_x\rangle, \dots, |\uparrow_y \text{ or } \downarrow_y\rangle, \dots, |\uparrow_z \text{ or } \downarrow_z\rangle, \dots$. We have an infinite number of hidden variables to represent all directions of measurement. Principally, by choosing appropriate $\rho(\vec{\lambda})$ one can always fulfill QS-I and II. For example, the following scheme gives the correct results on measurement of σ_r . For direction $\vec{r} = (r_x, r_y, r_z)$, we choose $\lambda_{\vec{r}} \in \{-\frac{1}{2}, \frac{1}{2}\}$, a two-value discrete random variable as the hidden variable. Then we require the partial trace except direction \vec{r} of ρ^{hvt} gives,

$$\rho_{\vec{r}}^{hvt} \triangleq \text{tr}^{-\vec{r}}(\rho^{hvt}) = \frac{1+r_x}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{1-r_x}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}, \quad (40)$$

for example, by requiring

$$\rho^{hvt} = \Pi_{\vec{r}} \otimes \rho_{\vec{r}}^{hvt}. \quad (41)$$

Notice the product state in equ(41) is just one example, not necessary required, while equ(40) is a strict requirement. There are many more density matrices in the form of equ(39) and satisfying equ(40). Independence of pure elementary events does not lead to independent product states.

One can check this satisfies QS-I and II for measurement on an arbitrary \vec{r} direction. The outcomes could be \uparrow or \downarrow with probability of $\frac{1+r_x}{2}$ and $\frac{1-r_x}{2}$ respectively. Furthermore, this HVT does not require contextuality between object and observer. After the partial trace only the desired direction will survive. The partial trace is a standard procedure for independent random variable. The above explicitly constructed density matrix gives the correct results for measurement on any directions. We see that our HVT is at least as valid as Bell's HVT on a spin $\frac{1}{2}$ object[11] as they both respect QS-I and QS-II. It is less controversial than our former exclusive-event HVT.

Will this HVT realize QS-III and QS-IV? The answer is "yes" for QS-III. According to CPT-III, after measurement, the system stays at the state observed for the observable and all the others remain at the same states. For example, when we measure σ_z with the outcome being *up*, the state after measurement is

$$\rho_{after}^{hvt} = \frac{|\uparrow\rangle_z \langle\uparrow|_z \otimes \langle\uparrow|_z \rho_{before}^{hvt} |\uparrow\rangle_z}{tr^{-z} \left(\langle\uparrow|_z \rho_{before}^{hvt} |\uparrow\rangle_z \right)}. \quad (42)$$

If measured on σ_z again the outcome is still *up*. What if the second measurement is on a different direction, say σ_x ? We have,

$$\begin{aligned} p_{z_{up}, x_{down}} &= tr \left(|\downarrow\rangle_x \langle\downarrow|_x \rho_{after}^{hvt} \right) \\ &= \frac{tr^{-x, -z} \left(\langle\uparrow|_z \langle\downarrow|_x \rho_{before}^{hvt} |\downarrow\rangle_x |\uparrow\rangle_z \right)}{tr^{-z} \left(\langle\uparrow|_z \rho_{before}^{hvt} |\uparrow\rangle_z \right)} \\ &= 0, \end{aligned} \quad (43)$$

where we make use of $\langle\downarrow|_x \rho_{before}^{hvt} |\downarrow\rangle_x = 0$, the fact that the state is initially *x* direction *up*. However, this number is expected to be $\frac{1}{2}$. This shows our HVT does not respect QS-IV if CPT-III holds. Thus we need to modify CPT-III to the following,

CPT-III'' After a measurement, the object stays at the state which guarantees a repeat measurement gives the right result stated in QS-III and QS-IV. For example, for spin- $\frac{1}{2}$ after a measurement on $\sigma_{\vec{r}_0}$ and a *up*-state is recorded, it stays at, ρ^{hvt} satisfying $\rho_{\vec{r}}^{hvt} = tr^{-\vec{r}}(\rho^{hvt})$,

$$\rho_{\vec{r}}^{hvt} = \frac{1 + \vec{r} \cdot \vec{r}_0}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{1 - \vec{r} \cdot \vec{r}_0}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}, \quad (44)$$

If a *down*-state is recorded after measurement of $\sigma_{\vec{r}_0}$, we can simply replace \vec{r}_0 with $-\vec{r}_0$ in equ(44).

This version of CPT-III has the same inconsistency with the last exclusive-event HVT. Given a *x* direction *up*-state, represented by equ(40), if we want to calculate the probability of *x* direction *up*-state, we do

$$p_{x_{up}} = tr \left(|\uparrow\rangle_x \langle\uparrow|_x \rho^{hvt} \right) = 1. \quad (45)$$

This gives us the correct result, however, with the assumption that $|\uparrow\rangle_x \langle\uparrow|_x$ refers to the *x* direction *up*-state, which is not the state which really means the *x* direction *up*-state as in equ(40).

C. Final HVT: equivalent class on the set of density matrices

We have noticed that the density matrix in product form as in equ(41) is just a special case of equ(39). Due to equ(40), all qualified density matrices give correct results on measurement of $\sigma_{\vec{r}}$ along an arbitrary direction \vec{r} . One may tell the difference among them if measurements along different directions are performed simultaneously. Either because reality forbids us to do so, or because we do not have the technology yet, we are not able to perform such measurements. Before the possibility of those simultaneous measurements can be resolved, we do not know principally which specific one out of the set of general form is the right description of QS.

For now, we will focus on measurement along one direction. For this case, our density matrix based description of quantum system is redundant. An equivalent class over the whole set of density matrices in the general form of equ(39) can be defined as following: Two density matrices ρ^a, ρ^b are regarded as equivalent if and only if they lead to the same reduced density matrix,

$$tr^{-\vec{r}}(\rho^a) = tr^{-\vec{r}}(\rho^b), \forall \vec{r} \in \mathbb{R}^3. \quad (46)$$

Our state is represented by the those equivalent class as required by QS-V.

For example, the first preparation gives us state,

$$\rho^I = \frac{1}{4}\rho^{11} + \frac{3}{4}\rho^{12}, \quad (47)$$

ρ^{11} is a product state of $\rho_{\vec{r}}^{11}$,

$$\rho_{\vec{r}}^{11} = \frac{1+r_z}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{1-r_z}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}. \quad (48)$$

ρ^{12} is a product state of $\rho_{\vec{r}}^{12}$,

$$\rho_{\vec{r}}^{12} = \frac{1-r_z}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{1+r_z}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}. \quad (49)$$

The second preparation gives us state,

$$\rho^{II} = \frac{1}{2}\rho^{21} + \frac{1}{2}\rho^{22}, \quad (50)$$

ρ^{21} is a product state of $\rho_{\vec{r}}^{21}$,

$$\rho_{\vec{r}}^{21} = \frac{1 - r_y \frac{\sqrt{3}}{2} - r_z \frac{1}{2}}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{1 + r_y \frac{\sqrt{3}}{2} + r_z \frac{1}{2}}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}. \quad (51)$$

ρ^{22} is a product state of $\rho_{\vec{r}}^{22}$,

$$\rho_{\vec{r}}^{22} = \frac{1 + r_y \frac{\sqrt{3}}{2} - r_z \frac{1}{2}}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} - \frac{1 + r_y \frac{\sqrt{3}}{2} + r_z \frac{1}{2}}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}. \quad (52)$$

In fact, we can also write down a product form density matrix according to equ(1). ρ^{III} is a product state of $\rho_{\vec{r}}^{III}$,

$$\rho_{\vec{r}}^{III} = \frac{2-r_z}{4} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{2+r_z}{4} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}} \quad (53)$$

It is straightforward to check that

$$\text{tr}^{-\vec{r}}(\rho^I) = \text{tr}^{-\vec{r}}(\rho^{II}) = \text{tr}^{-\vec{r}}(\rho^{III}). \quad (54)$$

But

$$\rho^I \neq \rho^{III} \neq \rho^{II}. \quad (55)$$

Without the equivalent class, they are different density matrices. Then our HVT does not respect QS-V. With it, they are regarded as the same so that QS-V is satisfied.

If in the future, we would be able to measure σ along several directions simultaneously, it would force us to pick up one specific form out of the whole set satisfying equ(40) and equ(39) and to discard the above equivalent class. For now, this is the final HVT we can propose as far as we require it to respect all five QS facts.

D. Test of all above HVTs against measurements

Besides our EHVT and IHVT, let us also check Bell's HVT and Bohm's HVT against our five QS facts. Imagine we are given one of the five states below and a measurement device as will be explained. Then we are asked to find out which one is the real state of the given object.

A A quantum spin- $\frac{1}{2}$ at $\rho_0 = |\uparrow\rangle_x \langle\uparrow|_x$.

B A classical two-face dice at state $\rho_0 = \frac{1}{2} |\uparrow\rangle_z \langle\uparrow|_z + \frac{1}{2} |\downarrow\rangle_z \langle\downarrow|_z$.

C A classical vector pointing to arbitrary directions with probability $\rho_0 = \frac{1}{4\pi} \iint d\theta d\phi \sin \theta \frac{1+\sin \theta \cos \phi}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}}$. If we rewrite state A in a spin coherent basis, we will get the same distribution. The only difference is that here in treating it like a classical object, we further assume the basis is orthogonal. It is a state in the form of an exclusive-event HVT (EHVT).

D A classical object at state $\rho_0 = \Pi_{\vec{r} \in \mathbb{D}} \otimes \rho_0^{\vec{r}}$, where $\rho_0^{\vec{r}} = \frac{1+\sin \theta \cos \phi}{2} |\uparrow\rangle_{\vec{r}} \langle\uparrow|_{\vec{r}} + \frac{1-\sin \theta \cos \phi}{2} |\downarrow\rangle_{\vec{r}} \langle\downarrow|_{\vec{r}}$. Here $\mathbb{D} = ([0, \frac{\pi}{2}] \otimes [0, 2\pi)) \cup (\{\frac{\pi}{2}\} \otimes [0, \pi))$, which denotes half of all direction vector \vec{r} . This is a state in the form of an independent-event-equivalent-class HVT (IHVT).

E Bell's hidden variable theory of spin- $\frac{1}{2}$ object[11]. Hidden variable $\lambda \in [-\frac{1}{2}, \frac{1}{2}]$ uniformly distributed. Given a specific λ , measurement on Pauli matrix $\vec{\beta} \cdot \vec{\sigma}$ on direction $\vec{\beta}$ yields, $\text{sign}(\lambda + \frac{1}{2}\beta_x) \text{sign}(X)$, where $X = \beta_x$ if $\beta_x \neq 0$, $X = \beta_y$ if $\beta_y \neq 0, \beta_x = 0$ and $X = \beta_z$ if $\beta_z \neq 0, \beta_x = 0, \beta_y = 0$. Here we changed the expression accordingly to represent the x direction up state.

The measurement device has an indicator showing a positive/negative value if the object is along the same/opposite direction. One can control the direction of the device. When its direction is not parallel or opposite to the object's direction, it will not be activated. Assume this device is sharp so that it will not respond to even a slight mis-matching. The device works on both classical and quantum systems.

Define the activation ratio Q as the ratio between times when the device is activated out of the total times the device is used, and define the up -state probability P as the ratio between the numern of positive values out of the times when the device is activated. We want to check if the above five states give us different values of Q and P during measurements. First, assume the device is along the z direction. We see from Table I that from the values

TABLE I: Values of Q and P with device along the z direction

	A	B	C	D	E
Q	1	1	$\ll 1$ [16]	1	1
P	0.5	0.5	0.5	0.5	0.5

of Q , state C is different from state A .

Next we adjust the device to the x direction. We see

TABLE II: Values of Q and P with device along the x direction

	A	B	C	D	E
Q	1	0	$\ll 1$	1	1
P	1	NA	1	1	1

from Table II that from the values of Q state B is different from state A . However, those measurements do not differentiate state A , D and E . The fact those two states D and E both respects QS-I and QS-II, makes them very good counterexamples of Von Neumann's proof of impossibility of HVT[3]. This is exactly made possible by that operators in those two theories do not obey (27) the linear relation between operators even when operators' averages have those linear relation. Such relation between operators is too restrictively assumed in Von Neumann's proof and leads to impossibility[3, 12].

In order to differentiate state A , D and E , we have to perform repeat measurement, say first along the z direction and then along the x direction. In dealing with repeat measurement, we need some rules to determine the object's state right after the first measurement. Here we first assume both CPT-III and QM-III hold. In the following table, we list only values of Q and P after the second measurement. From Table III the values of P there we find that state D and E are different from state A . That is we can distinguish a quantum state with Bell's HVT and our IHVT state by measurements if

TABLE III: Values of Q and P during the second measurement in a repeat measurement with device along the z and then the x direction, assuming both CPT-III and QM-III hold

	A	B	C	D	E
Q_2	1	0	0	1	1
P_2	0.5	NA	NA	1	1

CPT-III/QM-III holds. As for state D , this can be seen from,

$$\rho_1 = \frac{|\uparrow\rangle_z \langle\uparrow|_z \otimes \langle\uparrow|_z \rho_0 |\uparrow\rangle_z}{\text{tr}^{-z}(\langle\uparrow|_z \rho_0 |\uparrow\rangle_z)}, \quad (56)$$

and

$$P_2 = \frac{\text{tr}^{-x,-z}(\langle\uparrow|_z \langle\uparrow|_x \rho_0 |\uparrow\rangle_x |\uparrow\rangle_z)}{\text{tr}^{-z}(\langle\uparrow|_z \rho_0 |\uparrow\rangle_z)} = 1. \quad (57)$$

As for state E , let's assume $\lambda = \lambda^*$ after the first measurement, then for the second measurement one will get,

$$\text{sign}\left(\lambda^* + \frac{1}{2}\beta_x\right) \text{sign}(\beta_x) = \text{sign}\left(\lambda^* + \frac{1}{2}\right) = 1, \forall \lambda^*. \quad (58)$$

If we are allowed to relax CPT-III then it is always possible to adjust ρ_1 for state D and adjust the proposed measurement result for state E after the first measurement to make $P_2 = 0.5$. We have done so for state D in CPT-III''. And here we can adjust state E to satisfy the requirement. That is if we get the *up/down*-state in the first measurement, for arbitrary second measurement of $\vec{\beta} \cdot \vec{\sigma}$

$$\text{sign}\left(\lambda \pm \frac{1}{2}\beta_z\right) \text{sign}(X), \quad (59)$$

where

$$X = \begin{cases} \beta_z & \text{if } \beta_z \neq 0 \\ \beta_x & \text{if } \beta_z = 0, \beta_x \neq 0 \\ \beta_y & \text{if } \beta_z = 0, \beta_x = 0, \beta_y \neq 0 \end{cases}. \quad (60)$$

In that case, state D and E are indistinguishable from state A under all measurements, while state D and E are classical states and state A is a quantum state, as we see in Table IV.

TABLE IV: Values of Q and P during the second measurement in a repeat measurement with device along z and then x direction, with CPT-III adjusted accordingly

	A	B	C	D	E
Q_2	1	0	$\ll 1$	1	1
P_2	0.5	NA	0.5	0.5	0.5

From above comparison, we see that when repeat measurement is taken into consideration and CPT-III holds, none of all five theories respects all five QS facts. Only when we relax CPT-III, both Bell's HVT and our IHVT provide alternative theory for quantum systems. Unlike Bell's HVT theory, in our IHVT state, we have explicitly written down the state in a density matrix form, so it can be generalized for any objects not only spin- $\frac{1}{2}$ particles. In this sense this work can be seen as a development of Bell's HVT. Another thing we would like to point out is the relation between our IHVT and Bohm's HVT[12]: in the following sense, our IHVT provides exactly the explicit form of a state of Bohm's HVT.

Originally Bohm's HVT gave only a classical HVT based interpretation of measurement process on a single direction. Since we are free to choose an arbitrary direction, we need to generalize the theory a little bit. Basically it says during measurement process of a specific direction \vec{r} , system evolves according to the following equation system,

$$\begin{cases} \frac{dJ_{\vec{r}}^1}{dt} = 2\gamma(R^1 - R^2) J_{\vec{r}}^1 J_{\vec{r}}^2 \\ \frac{dJ_{\vec{r}}^2}{dt} = 2\gamma(R^2 - R^1) J_{\vec{r}}^2 J_{\vec{r}}^1 \end{cases}, \quad (61)$$

where $R^i = \frac{|J_{\vec{r}}^i|^2}{|\xi_{\vec{r}}^i|^2}$ and $\xi_{\vec{r}}^i$ are those hidden variables. Instead of quantum wavefunction ψ here we take $J_{\vec{r}}^i$ as our fundamental variables since only $J_{\vec{r}}^1 = |\langle\uparrow_{\vec{r}}|\psi\rangle|^2$ and $J_{\vec{r}}^2 = |\langle\downarrow_{\vec{r}}|\psi\rangle|^2$ is used in those equations. Then if we only focus on state representing this direction only, it can be written down as

$$\rho_{\vec{r}} = J_{\vec{r}}^1 |\uparrow_{\vec{r}}\rangle\langle\uparrow_{\vec{r}}| + J_{\vec{r}}^2 |\downarrow_{\vec{r}}\rangle\langle\downarrow_{\vec{r}}|. \quad (62)$$

From this point of view, (61) provides an explanation of the process that the above state in (62) turns into $|\uparrow_{\vec{r}}\rangle\langle\uparrow_{\vec{r}}|$ or $|\downarrow_{\vec{r}}\rangle\langle\downarrow_{\vec{r}}|$ at probability respectively $J_{\vec{r}}^1$ or $J_{\vec{r}}^2$. Now let's consider a separate measurement along another direction \vec{r}' . One possible way is to start from the quantum wavefunction ψ again to calculate $J_{\vec{r}'}^i$ and redo above procedure. This is actually not so bad but in this way this theory never gets rid of the quantum wavefunction. This makes Bohm's HVT only an alternative theory of quantum measurements but not a coherent theory covering both evolutions and measurements. There is however another way to recover the right prediction of measurement on \vec{r}' and it gets ride of quantum wavefunction totally. That is to assume that the HVT state is in fact,

$$\rho = \Pi_{\vec{r}} \otimes \rho_{\vec{r}}, \quad (63)$$

while $\rho_{\vec{r}}$ is given by (62) for a specific direction \vec{r} with proper predefined $J_{\vec{r}}^i$. We see that this is exactly state D , our IHVT state. We have shown that this state agrees with quantum mechanical state A on everything with CPT-III replaced by CPT-III''.

However, this IHVT is far from a standard CPT. To summarize, IHVT satisfies QS-I and QS-II easily but

CPT-III needs to be modified to make it satisfy QS-III and QS-IV. IHVT does not require contextuality between object and observer as EHVT does. But both suffer from the same inconsistency problem: two different expressions are used to represent the same state for two different purposes. Furthermore, both discard inherent relation among operators as in equ(27) by treating operators independently or exclusively. We find that all of the above has made HVT less understandable than the usual QM, which has none of above problems and respects all five QS facts. Therefore, we would like to conclude that we have ruled out HVT just from theoretical consideration and just by considering a spin- $\frac{1}{2}$ object.

If one is still willing to pay all the prices mentioned above, then we are also willing to go a little further to show that this IHVT conceals something else which one may not want in a theory of physics. We will apply this theory onto the description of singlet state, the entangled state used in the discussion of Bell's inequality.

VI. CPT FOR TWO-SPIN SYSTEM

The quantum density matrix form of a singlet state is

$$\rho^q = \frac{1}{2} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)(\langle\uparrow\downarrow| - \langle\downarrow\uparrow|), \quad (64)$$

where $|\uparrow\downarrow\rangle$ can be regarded as eigenstates on an arbitrary direction. The correlated quantum measurement of the \vec{r}_1 -direction on spin 1 and \vec{r}_2 on spin 2 gives

$$\langle\sigma_{\vec{r}_1}\sigma_{\vec{r}_2}\rangle = -\vec{r}_1 \cdot \vec{r}_2 \text{ and } \sigma_{\vec{r}_1}\sigma_{\vec{r}_2} = \pm 1. \quad (65)$$

A measurement on a single spin along any direction gives

$$\langle\sigma_{\vec{r}}\rangle = 0 \text{ and } \sigma_{\vec{r}} = \pm 1. \quad (66)$$

A successful HVT theory should give the two above results. Besides, for a repeat measurement, HVT should also give the correct results depending on the outcome from the first measurement. Although Bell's inequality has generally proved that through local classical theory it is impossible to achieve this, here, we will construct one such state, in the form of a classical density matrix, that does in fact achieve this. We will find out the cost of such a theory.

To denote a state in IHVT, one example of an equivalent class is used to represent the whole class. We should check if the following state respects all the QS facts. A reduced density matrix for two spins on \hat{z} and \hat{r} is

$$\rho_{\vec{r}_1, \vec{r}_2} = \frac{1-\vec{r}_1 \cdot \vec{r}_2}{4} (|\uparrow_{\vec{r}_1}\uparrow_{\vec{r}_2}\rangle\langle\uparrow_{\vec{r}_1}\uparrow_{\vec{r}_2}| + |\downarrow_{\vec{r}_1}\downarrow_{\vec{r}_2}\rangle\langle\downarrow_{\vec{r}_1}\downarrow_{\vec{r}_2}|) + \frac{1+\vec{r}_1 \cdot \vec{r}_2}{4} (|\uparrow_{\vec{r}_1}\downarrow_{\vec{r}_2}\rangle\langle\uparrow_{\vec{r}_1}\downarrow_{\vec{r}_2}| + |\downarrow_{\vec{r}_1}\uparrow_{\vec{r}_2}\rangle\langle\downarrow_{\vec{r}_1}\uparrow_{\vec{r}_2}|). \quad (67)$$

The whole density matrix is

$$\rho^{hvt} = \prod_{\vec{r}_1 \vec{r}_2} \otimes \rho_{\vec{r}_1 \vec{r}_2}, \quad (68)$$

The reduced density matrix satisfies equ(65) and equ(66), which is the content of QS-I and QS-II. For QS-III and QS-IV, although we will skip the details here, a state after measurement can be constructed easily. Building a state on the equivalent classes solves the problem of QS-V. We have successfully constructed a classical theory for two-spin quantum system. It is a classical theory but it violates Bell's inequality. As we argued above, we already know that, due to inconvenience and inconsistency, this theory should not be preferred. However, we can still ask how can such a classical theory does succeed to give all expected results from QM? The answer is it includes non-local information.

In [13], Bell's inequality was proved more generally with only the locality assumption, their equ(2') uses

$$p_{1,2}(\lambda, a, b) = p_1(\lambda, a) p_2(\lambda, b), \quad (69)$$

where λ is a hidden variable independent of a, b to express the idea of measurement-independent reality of a quantum system. Since our IHVT violates Bell's inequality, we want to check if it respects the above equation. Consider the situation where we measure direction a and b on those two spins respectively.

$$\begin{aligned} \langle\hat{S}^1\hat{S}^2\rangle(a, b) &= \text{tr}(\hat{S}^1(a)\hat{S}^2(b)\rho(\vec{\lambda})) \\ &= \sum_{\lambda_{ab}} \langle\lambda_{ab}|\hat{S}^1(a)\hat{S}^2(b)|\lambda_{ab}\rangle f(\lambda_{ab}) \\ &= \sum_{\lambda_{ab}} s^1(a, \lambda_{ab}) s^2(b, \lambda_{ab}) f(\lambda_{ab}) \end{aligned} \quad (70)$$

The left hand side can be regarded as

$$\langle\hat{S}^1\hat{S}^2\rangle(a, b) = \sum_{\lambda_{ab}} s^1 s^2(a, b, \lambda_{ab}) f(\lambda_{ab}). \quad (71)$$

From the core of the integral, we see that

$$s^1 s^2(a, b, \lambda_{ab}) = s^1(a, \lambda_{ab}) s^2(b, \lambda_{ab}), \quad (72)$$

or generally,

$$s^1 s^2(a, b, \vec{\lambda}) = s^1(a, \vec{\lambda}) s^2(b, \vec{\lambda}). \quad (73)$$

Compared with equ(69), equ(73) does look like an expression of locality, with the difference that a single hidden variable is replaced by many hidden variables. However, it is this replacement that introduces non-local information, because the effective one out of $\vec{\lambda}$ is λ_{ab} , which does depend on both a and b , the measurements on both spins. During the measurement process, a sample should be drawn from an effective probability distribution. And the effective one has to be determined through information with both directions a and b together. It is as if the system has to know both directions to make its decision. It is definitely contextual.

We have explicitly shown the place non-locality comes into QM. When the classical theory is used to describe QM, we have to require non-local information. If this non-locality is unacceptable, then we should rule out the idea of HVT. However, this never means QM in its own language requires non-local information. This is a topic which has never been addressed in this paper.

VII. CONCLUSION AND DISCUSSION

In a summary, to find a classical theory respecting all five QS facts, our conclusion is: first, single variable HVT is incompatible with non-commutative relation between operators; second, even if all operators are commutative, the inherent relation between them has to be abandoned; third, the exclusive-event HVT requires contextuality between object and observer; fourth, both EHVT and IHVT suffer from the inconsistency problem: the expression used to denote the state is different with the one used to recover probability; and at last, IHVT is shown to imply non-locality. We find the price is unreasonably high: even after we accept the non-locality, CPT-III need to be twisted. And due to those twists, such a classical system could no longer be broadcasted. Noticing CPT-III is essential to make it possible to broadcast a classical system. The possibility of being broadcasted is one key fact in understanding of classical measurement. Even theoretically, not depending on the experimental test of Bell's inequality, the idea of HVT should be discarded from theory of quantum systems.

In another words, under reasonable consideration it is impossible to map a full-structure density matrix to a diagonal density matrix. With this conclusion in mind, we may say that although the current language of QM may not be the ultimate one, any equivalent language should include existence of off-diagonal elements of the density matrix and allow vectors to be transformed from one basis to another, which is only possible when operators do

not always commute with each other. We know quantum measurement is not equal to classical measurement of a TRCO. Classical measurement creates a broadcast, but quantum measurement does not.

Finally, we are not saying those are all the possibilities of CPT for QS. From the C^* -algebra point of view, what we have tried here are just two examples of multiplications between operators, $\sigma_{\vec{r}_1}\sigma_{\vec{r}_2} = 0$ for the exclusive case and $\sigma_{\vec{r}_1} \otimes \sigma_{\vec{r}_2}$ for the independent case. There could be some other kinds of algebras among operators. If we assume symmetry among σ operator on all directions, then those two are the only choices. Besides our own HVTs we have also examined Bell's HVT and Bohm's HVT and ruled them out based on repeat measurement and validity of CPT-III.

VIII. ACKNOWLEDGMENT

Thanks should first be given to Dr. Leslie Ballentine and Dr. Ian Affleck for their inspiring courses on quantum mechanics, giving a detailed and insightful check of the basis of quantum mechanics, and an fascinating introduction into the fundamental problems in quantum mechanics. Thanks should also be given to Dr. Mona Berciu for her comments and discussions on this work, and Dr. Robert Peter and Janelle von Dongen for reading through the manuscript. At last, we acknowledge the anonymous referee gratefully for the comments and suggestions during the revision.

-
- [1] D. Bohm, Phys. Rev. **85**, 166(1952).
 - [2] A. Gut, *Probability: A Graduate Course* (Springer-Verlag, 2005).
 - [3] Neumann J. von, *Mathematical Foundations of Quantum Mechanics* (Princeton, Princeton Univ. Press, 1955).
 - [4] A. Lupascu, S. Saito, T. Picot, P. C. De Groot, C. J. P. M. Harmans, J. E. Mooij, **3**, 119(2007).
 - [5] L. Ballentine, *Quantum Mechanics: A Modern Development* (World Scientific, 1998).
 - [6] William Arveson, *An Invitation to C^* -Algebra* (Springer-Verlag, 1981).
 - [7] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
 - [8] W. K. Wootters and W. H. Zurek, Nature, **299**, 802 (1982).
 - [9] H. Barnum, C.M. Caves, C.A. Fuchs, R. Jozsa and B. Schumacher, Phys. Rev. Lett, **76**2818(1996).
 - [10] J.S. Bell, Physics **1**, 195(1964).
 - [11] J.S. Bell, Rev. Mod. Phys. **38**, 447(1966).
 - [12] D. Bohm and J. Bub, Rev. Mod. Phys. **38**, 453(1966).
 - [13] J.F. Clauser and M.A. Horne, Phys. Rev. **D10**, 526(1974).
 - [14] A. Aspect, Nature **398**, 189(1999).
 - [15] M.O. Scully, Phys. Rev. D, **28**, 2477(1983).
 - [16] If we allow contextual communication between object and observer then Q could be equal to 1. Or Q could also be equal to 1 if our device is totally unsharp, say it reads positive when $\vec{r} \cdot \vec{r}_0 > 0$ where \vec{r} is the vector's direction and \vec{r}_0 is device's direction. But the later is not the case we are discussing and theoretically it is always possible to make the device sharp.